Modelli 1 @ Clamfim

info sul corso

Differenziabilità, Massimi e minimi liberi

20 settembre 2016

professor Daniele Ritelli

daniele.ritelli@unibo.it

Partial derivatives

The most natural way to define derivatives of functions of several variables is to allow one variable to move at a time.

If $f : \mathbb{R}^n \to \mathbb{R}$ is a real function of n variables the partial derivative f_{x_j} exists at a point $\boldsymbol{a}$ if and only if the limit

$$\frac{\partial f}{\partial x_j}(\boldsymbol{a}) := \lim_{h \to 0} \frac{f(\boldsymbol{a} + h\boldsymbol{e}_j) - f(\boldsymbol{a})}{h}$$

exists

Higher-order partial derivatives are defined by iteration. For example, the second-order partial derivative of f with respect to x_j and x_k is defined by

$$f_{x_j x_k} := \frac{\partial^2 f}{\partial x_k \partial x_j} := \frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_j}\right)$$

when it exists. Second-order partial derivatives are called mixed when $j \neq k$

Definition. Let V be a nonempty, open subset of $\mathbb{R}^n$, let $f : V \to \mathbb{R}$, and let $p \in \mathbb{N}$.

(i) f is said to be $\mathcal{C}^p$ on V if and only if each partial derivative of f of order $k \leq p$ exists and is continuous on V.

Definition. Let V be a nonempty, open subset of $\mathbb{R}^n$, let $f : V \to \mathbb{R}$, and let $p \in \mathbb{N}$.

(i) f is said to be $\mathcal{C}^p$ on V if and only if each partial derivative of f of order $k \leq p$ exists and is continuous on V.

(ii) f is said to be $\mathcal{C}^\infty$ on V if and only if f is $\mathcal{C}^p$ for all $p \in \mathbb{N}$

Definition. Let V be a nonempty, open subset of $\mathbb{R}^n$, let $f : V \to \mathbb{R}$, and let $p \in \mathbb{N}$.

(i) f is said to be $\mathcal{C}^p$ on V if and only if each partial derivative of f of order $k \leq p$ exists and is continuous on V.

(ii) f is said to be $\mathcal{C}^\infty$ on V if and only if f is $\mathcal{C}^p$ for all $p \in \mathbb{N}$

$\mathcal{C}^p(V)$ denotes the set of functions that are $\mathcal{C}^p$ on an open set V

Theorem. [Clairaut-Schwarz] Suppose that V is open in $\mathbb{R}^2$, that $(a, b) \in V$, and that $f : V \to \mathbb{R}$. If f is $\mathcal{C}^2$ on V

$$\frac{\partial^2 f}{\partial y \partial x}(a, b) = \frac{\partial^2 f}{\partial x \partial y}(a, b)$$

Theorem. [Clairaut-Schwarz] Suppose that V is open in $\mathbb{R}^2$, that $(a, b) \in V$, and that $f : V \to \mathbb{R}$. If f is $\mathcal{C}^2$ on V

$$\frac{\partial^2 f}{\partial y \partial x}(a, b) = \frac{\partial^2 f}{\partial x \partial y}(a, b)$$

Theorem. If f is $\mathcal{C}^2$ on an open subset V of $\mathbb{R}^n$, if $\boldsymbol{a} \in V$, and if $j \neq k$, then

$$\frac{\partial^2 f}{\partial x_j \partial x_k}(\boldsymbol{a}) = \frac{\partial^2 f}{\partial x_k \partial x_j}(\boldsymbol{a})$$

Differentiability. We define what it means for a vector function to be differentiable at a point. Whatever our definition, we expect two things: If f is differentiable at $\boldsymbol{a}$ then

(1) f will be continuous at $\boldsymbol{a}$

(2) all first-order partial derivatives of f will exist at $\boldsymbol{a}$.

Definition. Let f be a real function of n variables. f is said to be differentiable at a point $\boldsymbol{a} \in \mathbb{R}^n$ if and only if there is an open set V containing $\boldsymbol{a}$ such that $f : V \to \mathbb{R}$ and there is a $\boldsymbol{d} \in \mathbb{R}^n$ such that

$$\lim_{\boldsymbol{h} \to 0} \frac{f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}) - \boldsymbol{d} \cdot \boldsymbol{h}}{||\boldsymbol{h}||} = 0$$

Definition. Let f be a real function of n variables. f is said to be differentiable at a point $\boldsymbol{a} \in \mathbb{R}^n$ if and only if there is an open set V containing $\boldsymbol{a}$ such that $f : V \to \mathbb{R}$ and there is a $\boldsymbol{d} \in \mathbb{R}^n$ such that

$$\lim_{\boldsymbol{h} \to 0} \frac{f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}) - \boldsymbol{d} \cdot \boldsymbol{h}}{||\boldsymbol{h}||} = 0$$

$\boldsymbol{d}$ is called total derivative of f at $\boldsymbol{a}$

Definition. Let f be a real function of n variables. f is said to be differentiable at a point $\boldsymbol{a} \in \mathbb{R}^n$ if and only if there is an open set V containing $\boldsymbol{a}$ such that $f : V \to \mathbb{R}$ and there is a $\boldsymbol{d} \in \mathbb{R}^n$ such that

$$\lim_{\boldsymbol{h} \to 0} \frac{f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}) - \boldsymbol{d} \cdot \boldsymbol{h}}{||\boldsymbol{h}||} = 0$$

$\boldsymbol{d}$ is called total derivative of f at $\boldsymbol{a}$

Theorem. If f is differentiable at $\boldsymbol{a}$, then:

(i) f is continuous at $\boldsymbol{a}$.

(ii) all first-order partial derivatives of f exist at $\boldsymbol{a}$.

(iii) $\boldsymbol{d} = \nabla f(\boldsymbol{a}) := \left(\frac{\partial f}{\partial x_1}(\boldsymbol{a}), \ldots, \frac{\partial f}{\partial x_n}(\boldsymbol{a}) \right)$

$\nabla f(\boldsymbol{a})$ is called the gradient or nabla of f in $\boldsymbol{a}$

To reverse the statement of the last Theorem if we strengthen the conclusion, we can obtain a reverse implication.

Theorem. Let V be open in $\mathbb{R}^n$, let $\boldsymbol{a} \in V$, and suppose that $f : V \to \mathbb{R}$. If all first-order partial derivatives of f exist in V and are continuous at $\boldsymbol{a}$, then f is differentiable at $\boldsymbol{a}$.

Chain Rule Suppose that $g = (g_1, \ldots, g_n)$ is a vector function from $I \subseteq \mathbb{R}$ to $\mathbb{R}^n$, being I an open interval and $f : g(I) \subseteq \mathbb{R}^n \to \mathbb{R}$. If each of the components g_j of g is differentiable at $t_0 \in I$ and if f is differentiable at $\boldsymbol{a} = (g_1(t_0), \ldots, g_n(t_0))$ then $\varphi(t) := f(g(t))$ is differentiable at t_0 and

$$\varphi'(t_0) = \nabla f(\boldsymbol{a}) \cdot g'(t_0)$$

where we set

$$g'(t_0) := \big(g_1'(t_0), \ldots, g_n'(t_0)\big)$$

and $\cdot$ is the dot (inner) product in $\mathbb{R}^n$

Jacobian matrix

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ a function from Euclidean n-space to Euclidean m-space. Such a function is given by m real-valued component functions, $f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)$. The partial derivatives of all these functions (if they exist) can be organized in an m-by-n matrix, the Jacobian matrix J of f, as follows:

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} := \frac{\partial(f_1, \ldots f_m)}{\partial(x_1, \ldots, x_n)}$$

Jacobian matrix

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ a function from Euclidean n-space to Euclidean m-space. Such a function is given by m real-valued component functions, $f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)$. The partial derivatives of all these functions (if they exist) can be organized in an m-by-n matrix, the Jacobian matrix J of f, as follows:

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} := \frac{\partial(f_1, \ldots f_m)}{\partial(x_1, \ldots, x_n)}$$

The i^{th} row $(i = 1, \ldots, m)$ of this matrix corresponds to the gradient of the i^{th}component function ∇f_i

Exercise Given $u(x, y) = x^3 + y^3, \quad x,\, y \in \mathbb{R}$ show that u verifyies

$$\boldsymbol{x^2 u_{xx} + 2xy u_{xy} + y^2 u_{yy} = 6u}$$

Exercise Given $u(x,y) = x^3 + y^3, \quad x,\, y \in \mathbb{R}$ show that u verifyies

$$\boldsymbol{x^2 u_{xx} + 2xy u_{xy} + y^2 u_{yy} = 6u}$$

Exercise Given $u(x,y) = x^2 + xy + y^2$ show that u verifyies

$$\boldsymbol{u_{xx} + u_{xy} - u_{yy} = 1}$$

Exercise Given $u(x,y) = x^3 + y^3, \quad x,\, y \in \mathbb{R}$ show that u verifyies

$$\boldsymbol{x^2 u_{xx} + 2xy u_{xy} + y^2 u_{yy} = 6u}$$

Exercise Given $u(x,y) = x^2 + xy + y^2$ show that u verifyies

$$\boldsymbol{u_{xx} + u_{xy} - u_{yy} = 1}$$

Exercise Given $u(x,t) = e^{-t} \sin x, \quad t,\, x \in \mathbb{R}$ show that u verifies the heat equation: i.e. $\boldsymbol{u_t = u_{xx}}$

Exercise Given $u(x,y) = x^3 + y^3, \quad x,\, y \in \mathbb{R}$ show that u verifyies

$$\boldsymbol{x^2 u_{xx} + 2xyu_{xy} + y^2 u_{yy} = 6u}$$

Exercise Given $u(x,y) = x^2 + xy + y^2$ show that u verifyies

$$\boldsymbol{u_{xx} + u_{xy} - u_{yy} = 1}$$

Exercise Given $u(x,t) = e^{-t} \sin x, \quad t,\, x \in \mathbb{R}$ show that u verifies the heat equation: i.e. $\boldsymbol{u_t = u_{xx}}$

Exercise Show that $u(x,y) := f(xy)$ verifyies partial differential equation

$$\boldsymbol{x \frac{\partial u}{\partial x} - y \frac{\partial u}{\partial y} = 0}$$

Positive homogeneous functions

A function $f : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ is positive homogeneous of degree k if for any $\boldsymbol{x} \in: \mathbb{R}^n \setminus \{0\}$

$$f(\alpha \boldsymbol{x}) = \alpha^k f(\boldsymbol{x})$$

Positive homogeneous functions

A function $f : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ is positive homogeneous of degree k if for any $\boldsymbol{x} \in: \mathbb{R}^n \setminus \{0\}$

$$f(\alpha \boldsymbol{x}) = \alpha^k f(\boldsymbol{x})$$

Euler's homogeneous function theorem states that if $f : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ is continuously differentiable, then f is positive homogeneous of degree k if and only if

$$\boldsymbol{x} \cdot \nabla f(\boldsymbol{x}) = k f(\boldsymbol{x})$$

Given

$$f(x, y) = \sqrt{x^4 + y^4 + x^2 y^2}$$

establish if it is positive homogeneous and, in case, establish the degree of homogeneity. Then verify thesis of Euler theorem

Critical points

Definition. Let V be an open set in $\mathbb{R}^n$, let $\boldsymbol{a} \in V$ and suppose that $f : V \to \mathbb{R}$.

(i) $f(\boldsymbol{a})$ is called a local minimum of f if and only if there is an $r > 0$ such that $f(\boldsymbol{a}) \leq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in B_r(\boldsymbol{a})$

Critical points

Definition. Let V be an open set in $\mathbb{R}^n$, let $\boldsymbol{a} \in V$ and suppose that $f : V \to \mathbb{R}$.

(i) $f(\boldsymbol{a})$ is called a local minimum of f if and only if there is an $r > 0$ such that $f(\boldsymbol{a}) \leq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in B_r(\boldsymbol{a})$

(ii) $f(\boldsymbol{a})$ is called a local maximum of f if and only if there is an $r > 0$ such that $f(\boldsymbol{a}) \geq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in B_r(\boldsymbol{a})$

Critical points

Definition. Let V be an open set in $\mathbb{R}^n$, let $\boldsymbol{a} \in V$ and suppose that $f : V \to \mathbb{R}$.

(i) $f(\boldsymbol{a})$ is called a local minimum of f if and only if there is an $r > 0$ such that $f(\boldsymbol{a}) \leq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in B_r(\boldsymbol{a})$

(ii) $f(\boldsymbol{a})$ is called a local maximum of f if and only if there is an $r > 0$ such that $f(\boldsymbol{a}) \geq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in B_r(\boldsymbol{a})$

(iii) $f(\boldsymbol{a})$ is called a local extremum of f if and only if $f(\boldsymbol{a})$ is a local maximum or a local minimum of f.

Remark. If the first-order partial derivatives of f exist at $\boldsymbol{a}$, and $f(\boldsymbol{a})$ is a local extremum of f, then $\nabla f(\boldsymbol{a}) = \mathbf{0}$.

Remark. If the first-order partial derivatives of f exist at $\boldsymbol{a}$, and $f(\boldsymbol{a})$ is a local extremum of f, then $\nabla f(\boldsymbol{a}) = \mathbf{0}$.

In fact the one-dimensional function

$$g(t) = f(a_1, \ldots, a_{j-1}, t, a_{j+1}, \ldots, a_n)$$

has a local extremum at $t = a_j$ for each $j = 1, \ldots, n$. Hence, by the one-dimensional theory

$$\frac{\partial f}{\partial x_j}(\boldsymbol{a}) = g'(a_j) = 0$$

Remark. If the first-order partial derivatives of f exist at $\boldsymbol{a}$, and $f(\boldsymbol{a})$ is a local extremum of f, then $\nabla f(\boldsymbol{a}) = \mathbf{0}$.

In fact the one-dimensional function

$$g(t) = f(a_1, \dots, a_{j-1}, t, a_{j+1}, \dots, a_n)$$

has a local extremum at $t = a_j$ for each $j = 1, \dots, n$. Hence, by the one-dimensional theory

$$\frac{\partial f}{\partial x_j}(\boldsymbol{a}) = g'(a_j) = 0$$

As in the one-dimensional case, $\nabla f(\boldsymbol{a}) = \mathbf{0}$ is necessary but not sufficient for $f(\boldsymbol{a})$ to be a local extremum.

Remark. There exist continuously differentiable functions that satisfy $\nabla f(\boldsymbol{a}) = \mathbf{0}$ such that $f(\boldsymbol{a})$ is neither a local maximum nor a local minimum.

Consider for $n = 2$

$$f(x, y) = y^2 - x^2$$

Remark. There exist continuously differentiable functions that satisfy $\nabla f(\boldsymbol{a}) = \mathbf{0}$ such that $f(\boldsymbol{a})$ is neither a local maximum nor a local minimum.

Consider for $n = 2$

$$f(x, y) = y^2 - x^2$$

It is easy to check that $\nabla f(\mathbf{0}) = \mathbf{0}$ but the origin is a saddle point see figure

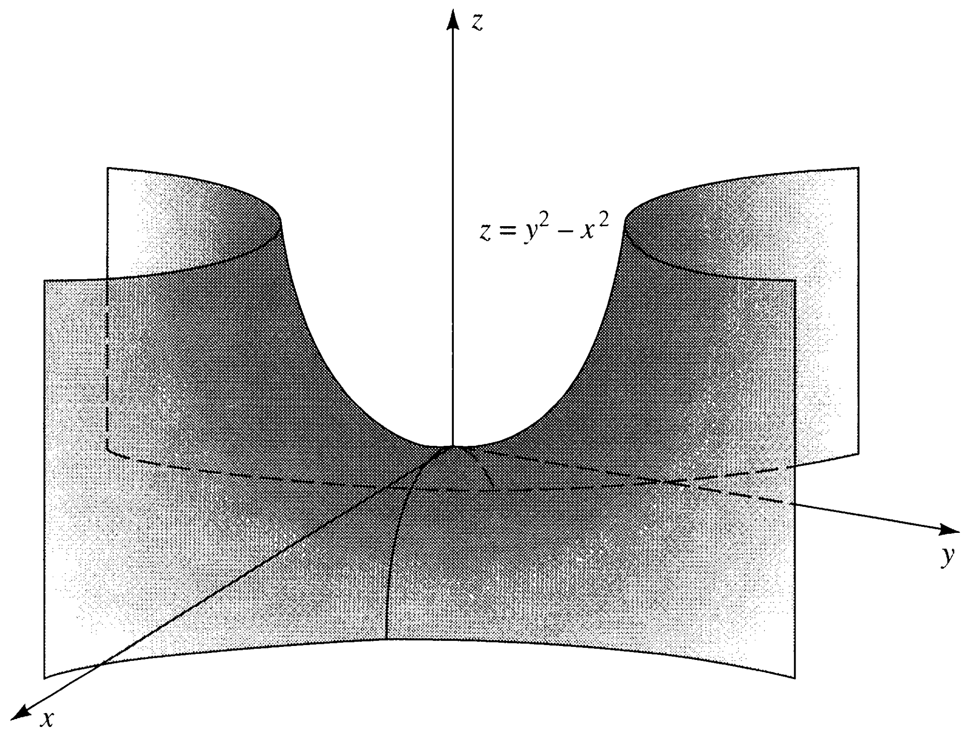

Figure 1: Saddle point

Definition. Let V be open in $\mathbb{R}^n$, let $\boldsymbol{a} \in V$, and let $f : V \to \mathbb{R}$ be differentiable at $\boldsymbol{a}$. Then $\boldsymbol{a}$ is called a saddle point of f if $\nabla f(\boldsymbol{a}) = \mathbf{0}$ and there is a $r_0 > 0$ such that given any $0 < \rho < r_0$ there are points $\boldsymbol{x}$, $\boldsymbol{y} \in B_\rho(\boldsymbol{a})$ that satisfy

$$f(\boldsymbol{x}) < f(\boldsymbol{a}) < f(\boldsymbol{y})$$

Hessian matrix. Let $V \subseteq \mathbb{R}^n$ an open set and let $f : V \to \mathbb{R}$ be a $\mathcal{C}^2$ function. The Hessian matrix of f at $\boldsymbol{x} \in V$ (or simply the Hessian) is the symmetric square matrix of second-order partial derivatives of f at $\boldsymbol{x}$:

$$H(f)(\boldsymbol{x}) := \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x})\right]_{i,j=1,\ldots,n}$$

Test for extrema and saddle points

Theorem. Let V be open in $\mathbb{R}^2$, $(a,b) \in V$, and suppose that $f : V \to \mathbb{R}$ satisfies $\nabla f(a,b) = \mathbf{0}$. Suppose further that $f \in \mathcal{C}^2$ and set

$$D := f_{xx}(a,b) f_{yy}(a,b) - f_{xy}^2(a,b)$$

Test for extrema and saddle points

Theorem. Let V be open in $\mathbb{R}^2$, $(a, b) \in V$, and suppose that $f : V \to \mathbb{R}$ satisfies $\nabla f(a, b) = \mathbf{0}$. Suppose further that $f \in \mathcal{C}^2$ and set

$$D := f_{xx}(a, b) f_{yy}(a, b) - f_{xy}^2(a, b)$$

(i) If $D > 0$ and $f_{xx}(a, b) > 0$ then $f(a, b)$ is a local minimum

Test for extrema and saddle points

Theorem. Let V be open in $\mathbb{R}^2$, $(a,b) \in V$, and suppose that $f : V \to \mathbb{R}$ satisfies $\nabla f(a,b) = \mathbf{0}$. Suppose further that $f \in \mathcal{C}^2$ and set

$$D := f_{xx}(a,b) f_{yy}(a,b) - f_{xy}^2(a,b)$$

(i) If $D > 0$ and $f_{xx}(a,b) > 0$ then $f(a,b)$ is a local minimum

(ii) If $D > 0$ and $f_{xx}(a,b) < 0$ then $f(a,b)$ is a local maximum

Test for extrema and saddle points

Theorem. Let V be open in $\mathbb{R}^2$, $(a, b) \in V$, and suppose that $f : V \to \mathbb{R}$ satisfies $\nabla f(a, b) = \mathbf{0}$. Suppose further that $f \in \mathcal{C}^2$ and set

$$D := f_{xx}(a, b) f_{yy}(a, b) - f_{xy}^2(a, b)$$

(i) If $D > 0$ and $f_{xx}(a, b) > 0$ then $f(a, b)$ is a local minimum

(ii) If $D > 0$ and $f_{xx}(a, b) < 0$ then $f(a, b)$ is a local maximum

(iii) If $D < 0$ the (a, b) is a saddle point

Test for extrema and saddle points

Theorem. Let V be open in $\mathbb{R}^2$, $(a,b) \in V$, and suppose that $f : V \to \mathbb{R}$ satisfies $\nabla f(a,b) = \mathbf{0}$. Suppose further that $f \in \mathcal{C}^2$ and set

$$D := f_{xx}(a,b) f_{yy}(a,b) - f_{xy}^2(a,b)$$

(i) If $D > 0$ and $f_{xx}(a,b) > 0$ then $f(a,b)$ is a local minimum

(ii) If $D > 0$ and $f_{xx}(a,b) < 0$ then $f(a,b)$ is a local maximum

(iii) If $D < 0$ the (a,b) is a saddle point

Observe that

$$D = \det[H(f)(a,b)]$$

Examples.

$$\boldsymbol{f(x,y) = x^3 + 6xy - 3y^2 + 2}$$

has a saddle point in $(a,b) = (0,0)$ and a local maximum in $(a,b) = (-2,-2)$

Examples.

$$\boldsymbol{f(x,y) = x^3 + 6xy - 3y^2 + 2}$$

has a saddle point in $(a,b) = (0,0)$ and a local maximum in $(a,b) = (-2,-2)$

$$\boldsymbol{f(x,y) = x^2 + y^3 - 2xy - y}$$

has a saddle point in $(a,b) = (-\frac{1}{3}, -\frac{1}{3})$ and a local minimum in $(a,b) = (1,1)$